A conjecture on the distribution of firm profit

Ian Wright[*][†]

Abstract

A common assumption of political economy is that profit rates across
firms or sectors tend to uniformity, and often models are formulated in
which this tendency is assumed to have been realised. But in reality this
tendency is never realised and the distribution of firm profits is not de-
generate but skewed to the right. The mode is less than the mean and
super-profits are present. To understand the distribution of firm prof-
its a general probabilistic argument is sketched that yields a candidate
functional form. The overall properties of the derived distribution are
qualitatively consistent with empirical measures, although there is more
work to be done.

Key words: firms, profit, economic, distribution, probabilistic

1 Introduction

Farjoun and Machover [2], dissatisfied with the concept of mechanical equilib-
rium applied to political economy and the concomitant assumption of a realised
uniform profit rate, outlined a probabilistic approach to political economy, which
replaced mechanical equilibrium with statistical equilibrium and a uniform profit
rate with a distribution of profit rates. They reasoned that the proportion of
industrial capital, out of the total capital invested in the economy, which finds
itself in any given profit bracket will be approximated by a gamma distribution,
by analogy with the distribution of kinetic energy in a gas at equilibrium. The
gamma distribution is a right-skewed distribution. They examined UK industry
data from 1972 and concluded that it was consistent with a gamma distribution.
Wells [6] examined the distributions of profit rates defined in a variety of ways
of over 100,000 UK firms and found right-skewness to be prevalent, but did
not investigate their functional form. Wright [7] measured the distribution of
firm profits in an agent-based model of a competitive economy, and found that
the distribution was right-skewed, although not well characterised by a gamma

[*]iKuni Inc., 3400 Hillview Avenue, Building 5, Palo Alto, CA 94304, USA. Email:
wrighti@acm.org, URL: ianusa.home.mindspring.com. Fax: +1 650 320 9827. Phone: +1
650 739 5355.

distribution, even when capital-weighted. Analysis of the model suggested that the profit distribution may be explained by general probabilistic laws.

The remainder of the paper outlines some theoretical assumptions and derives a candidate functional form for the distribution of firm profits.

2   A probabilistic argument

Under normal circumstances a firm expects that a worker adds a value to the product that is bound from below by the wage. A firm's markup on costs reflects this value expectation, which may or may not be validated in the market. Wages are normally paid in installments of between a week and one month, but the markup on costs is validated in the market at a frequency that depends on the rate at which a firm's goods and services are purchased by buyers. The frequency of payments to a firm differ widely and depend on the complexity of the product and the details of payment schedules (for example, compare a firm that sells sweets to a firm that sells battleships). The frequency mismatch between wage payments and revenue payments can be mitigated in many different ways, not least by the arrangement of capital loans. But whatever the frequency of sale or the complexity of the product a revenue payment to a firm partially reflects the value added by the firm's workers during a period of time. Assume that the revenue from the sale of a firm's product consists of a sum of market samples where each sample represents the value-added by a particular employee working for a small period of time, say an hour. Obviously, there are multiple and particular reasons why an individual worker adds more or less value to the firm's total product, most of which are difficult to measure, as partially reflected in the large variety of contested and negotiable compensation schemes. Although each worker normally adds value there is a great deal of local contingency. A worker may be a slacker or a workaholic, an easily replaceable administrator, or a unique, currently fashionable film star. Therefore, the precise value contribution of an individual worker to the product is highly complex and largely unknown, particularly when it is considered that the productive co-operation of many workers cannot be easily reduced to separate and orthogonal contributions, as is the case in highly creative industries with production processes that have yet to mature into separable, repeatable and well-defined tasks. This local contingency and indeterminacy is modelled by assuming that the value-added per worker-hour is a *random variable*. Consider that a worker i adds a monetary value, X_i, to a firm's product for every hour worked, where each X_i is an independent and identically distributed (iid) random variable, with mean μ_X and variance σ_X^2. The added value is assumed to be globally idd to reflect the common determinants of the value-creating power of an hour of work, but also random to model local contingencies. Negative X_i represents negative value-added, corresponding to cases in which the worker's labour reduces the value of inputs, for example the production of unwanted goods, or a slower than average work pace, and so forth.

Assume that the distribution of X_i is such that the Central Limit Theorem

(CLT) may be applied. Consider a single firm that sets in motion a total of n worker-hours during a single year. The firm's total value-added, S_n, may therefore be approximated by a normal distribution $S_n = \sum_{i=1}^{n} X_i \approx N(n\mu_X, n\sigma_X^2)$. The CLT approximation will improve with the size of the firm, but even for small firms the number of iid draws is large given the stated assumptions.

In reality the productivity of workers within firms is correlated. For example, employees of firms that employ state-of-the-art machinery, or are exceptionally well-organised, will all tend to add more value than employees of firms that employ out-of-date machinery or are badly organised. Although competitive processes tend to homogenise the value-added per worker, new innovations never cease, so that at any moment in time the employees of particular firm will be more or less productive than the average. A more accurate representation of value-added is obtained if each X_i is considered to be drawn from a distribution indexed by the firm that employs worker i, at the expense of a considerable increase in model complexity. However, the correlation of value-added within a large firm, which employs diverse skills and machinery to produce a variety of products, will be weak. Although a huge multinational is normally considered a single entity for the purpose of reporting profits, in reality it sets into a motion a large sample of different kinds of labours utilising different kinds of machinery and tools. Hence, for large firms the assumption that X_i is sampled from a single, economy-wide distribution is a reasonable approximation, for small firms less so. An advantage of modelling value-added per worker as a random variable is that it is possible that total value-added by a firm, S_n, is much higher or lower than the norm, but this event has low probability. The assumption of a single distribution that determines the value-added per worker is able to approximate the diverse productivities of individual firms.

Each worker costs a certain amount to employ during the year. This cost includes the wage, the cost of inputs used by the worker, the cost of wear and tear on any fixed capital, the cost of rent, local taxes and so forth, all of which may be differently reported due to local accountancy practices. Again, there is a great deal of contingency. Hence costs per worker-hour are also modelled as a random variable. Assume that a worker i costs a monetary value, Y_i, to productively employ per hour worked, where each Y_i is an idd random variable with mean μ_Y and variance σ_Y^2. This cost includes both the wage and capital costs per worker, and therefore effaces the distinction between variable and constant capital. Costs per worker-hour are also correlated at the firm level: the employees of different firms productively combine a greater or lesser amount of capital. A more accurate representation of costs would therefore consider the distribution of constant capital across firms conditional on local circumstances, such as firm size, but this extension is not pursued here. The assumption that cost per worker-hour is statistically unifrom across firms is an approximation, which, as for the case of value-added, improves with firm size, under the assumption of a tendency toward homogenisation due to competitive pressures.

Assume that the distribution of Y_i is such that the CLT may be applied. Hence a firm that sets in motion n worker-hours during a year has total costs that may be approximated by a normal distribution, $K_n = \sum_{i=1}^{n} Y_i \approx N(n\mu_Y, n\sigma_Y^2)$.

This approximation also improves with the size of the firm.

Different firms employee different numbers of workers and hence the amount of hours worked for each firm during a year will vary. Define the profit, P_n, of a firm that sets in motion n hours of labour in a single year as the ratio of value-added to costs, $P_n = S_n/K_n$, and assume that S_n and K_n are independent. P_n is the ratio of two normal variates. Its probability density function (pdf) may derived by the transformation method (or alternatively see [5]) to give:

$$
\begin{aligned}
f_{P_n}(p \mid n) &= \frac{\sqrt{n}\exp[-\frac{1}{2}n(\mu_X^2/\sigma_X^2 + \mu_Y^2/\sigma_Y^2)]}{4\pi(\sigma_X^2 + p^2\sigma_Y^2)^{3/2}} \\
&\quad \left(\frac{2}{\sqrt{n}}\sqrt{\lambda_1} + \sqrt{2\pi}\exp[\frac{n}{2}\frac{\lambda_2^2}{\lambda_1}]\lambda_2\left(1 + \Phi[\sqrt{\frac{n}{2}}\frac{\lambda_2}{\sqrt{\lambda_1}}]\right)\right)
\end{aligned} \quad (1)
$$

where

$$
\begin{aligned}
\lambda_1 &= \sigma_X^2\sigma_Y^2(\sigma_X^2 + p^2\sigma_Y^2) \\
\lambda_2 &= \mu_Y\sigma_X^2 + p\mu_X\sigma_Y^2 \\
\Phi(x) &= \frac{2}{\sqrt{\pi}}\int_0^x \exp^{-t^2} dt
\end{aligned}
$$

Equation (1) is the pdf of the rate-of-profit of a firm conditional on n, the number of hours worked for the firm per year.

Axtell [1] analysed US Census Bureau data for US firms trading between 1988 and 1997 and found that the firm size distribution, where size is measured by the number of employees, followed a special case of a power-law known as Zipf's law, and this relationship persisted from year to year despite the continual birth and demise of firms and other major economic changes. During this period the number of reported firms increased from 4.9 million to 5.5 million. Gaffeo et. al. [4] found that the size distribution of firms in the G7 group over the period 1987-2000 also followed a power-law, but only in limited cases was the power-law actually Zipf. Fujiwara et. al. [3] found that the Zipf law characterised the size distribution of about 260,000 large firms from 45 European countries during the years 1992–2001. A Zipf law implies that a majority of small firms coexist with a decreasing number of disproportionately large firms. Firm sizes theoretically range from 1 (a degenerate case of a self-employed worker) to the whole available workforce, representing a highly unlikely monopolisation of the whole economy by a single firm.

The empirical evidence implies that at any point in time the firm size distribution follows a power-law, and that this distribution is constant, despite the continual churning of firms in the economy (birth, death, shrinkage and growth). Firms hire and fire employees, and therefore the number of hours worked for a firm during a year depends on its particular historical growth pattern. To simplify, assume that the average number of employees per firm per year also follows a power-law. This approximation is reasonable if the growth trajectories of firms do not fluctuate too widely during the accounting period. Assume also

4

that every employee works the same number of hours in a year, which is a reasonable simplification. The firm hours per year is therefore a constant multiple of the number of firm employees. Firms with more employees proportionately set in motion more hours of labour. A constant multiple of a power-law variate is also a power-law variate. Hence, the firm size distribution has the same power-law form whether firm size is measured by employees or by the total number of hours worked by employees.

The unconditional rate-of-profit distribution can therefore be obtained by considering that the number of hours worked for a firm during a year is a random variable N distributed according to a Pareto (power-law) distribution:

$$f_N(n) = \frac{\alpha \beta^\alpha}{n^{\alpha+1}}$$

where α is the shape and β the location parameter. Assume that firm sizes range between m_1 hours, which represents a degenerate case of a self-employed worker who trades during the year, to m_2 hours, which represents a highly unlikely monopolisation of all social labour by a single huge firm ($m_2 >> m_1$). The truncated Pareto distribution

$$g_N(n) = f_N(n \mid m_1 < N \le m_2) = \frac{f_N(n)}{F_N(m_2) - F_N(m_1)} = \frac{n^{-(1+\alpha)} \alpha m_1^\alpha m_2^\alpha}{m_2^\alpha - m_1^\alpha}$$

where

$$f(n) = F'(n)$$

is formed to ensure that all the probability mass is between m_1 and m_2. Assume that m_2 is large so that the discrete firm size distribution can be approximated by the continuous distribution g_N.

By the Theorem of Total Probability the unconditional profit distribution $f_P(p)$ is given by:

$$f_P(p) = \int_{m_1}^{m_2} f_P(p \mid n) g_N(n) dn \tag{2}$$

Expression (2) defines the $g_N(n)$ parameter-mix of $f_P(p \mid N = n)$. The rate-of-profit variate is therefore composed of a parameter-mix of a ratio of independent normal variates each conditional on a firm size n, measured in hours per year, distributed according to a power-law. Writing (2) in full yields the pdf of firm profit:

$$
\begin{aligned}
f_P(p) &= \int_{m_1}^{m_2} \frac{\exp[-\frac{1}{2}n(\mu_X^2/\sigma_X^2 + \mu_Y^2/\sigma_Y^2)]}{4\pi(\sigma_X^2 + p^2\sigma_Y^2)^{3/2}} \\
&\quad \left(\frac{2}{\sqrt{n}}\sqrt{\lambda_1} + \sqrt{2\pi}\exp[\frac{n}{2}\frac{\lambda_2^2}{\lambda_1}]\lambda_2 \left(1 + \Phi[\sqrt{\frac{n}{2}}\frac{\lambda_2}{\sqrt{\lambda_1}}] \right) \right) \\
&\quad \frac{n^{-(\frac{1}{2}+\alpha)}\alpha m_1^\alpha m_2^\alpha}{m_2^\alpha - m_1^\alpha} \, dn
\end{aligned}
\tag{3}
$$

5

This distribution has 7 parameters: (i) μ_X, the mean value-added per worker-hour, (ii) σ_X^2, the variance of value-added per worker-hour, (iii) μ_Y, the mean cost per worker-hour, (iv) σ_Y^2, the variance of cost per worker-hour, (v) α, the Pareto exponent of the firm size power-law distribution, where size is measured in worker-hours per year, (vi) m_1, the number of hours worked by a single worker in a year, and (vii) m_2, the total number of hours worked in the whole economy during a year. Both percentage profit, $R = 100P$, and the growth rate of capital invested, $G = 1 + P$, are simple linear transforms of this distribution.

The parameters can be estimated from economic data and the resulting distribution compared to empirical rate-of-profit measures, under various simplifying assumptions about how profit is defined (e.g. see Wells [6]). A good fit would imply that the assumptions made in the theoretical derivation are empirically sound. Alternatively, best-fit parameters may be directly estimated from empirical data, for example by the method of maximum likelihood estimation, to determine how well the theoretical distribution can fit a set of empirical distributions. A good fit compared to other candidate functional forms would imply that a parameter-mix of a ratio of normal variates with parameters conditional on a power-law captures some essential structure of the determinants of firm profit, but it would not validate the theoretical derivation.
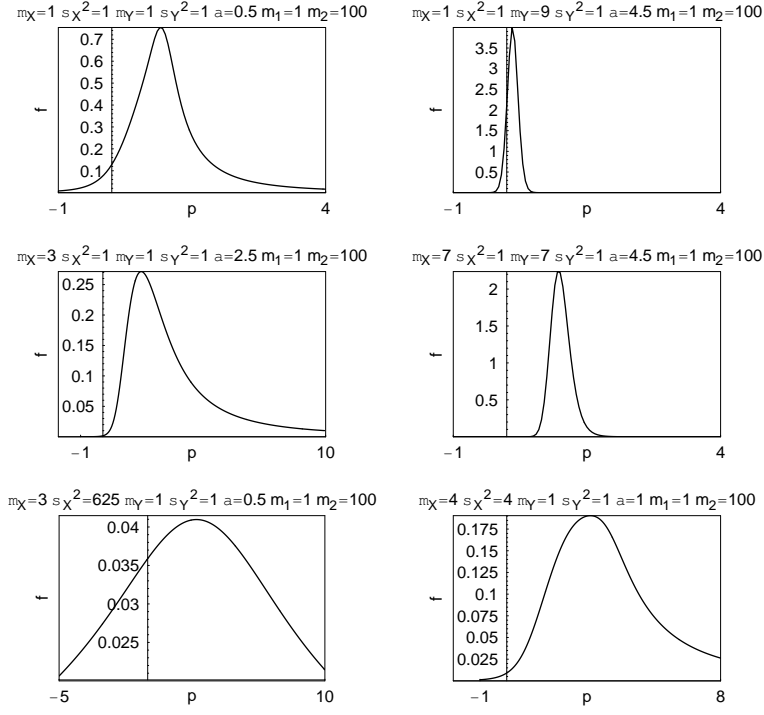
Equation (3) is difficult to analyse so numerical solutions are employed. Figure 1 graphs some representative numerical samples of the distribution. The samples range from sharply peaked symmetrical curves, in which most of the probability mass is concentrated about the mode, to less peaked distributions that are skewed to the right. Wells' [6] variety of profit measures yield distributions that share these characteristics, and therefore there is qualitative agreement between the theory and the empirical data. But clearly a full quantitative analysis is required.

Figure 2 graphs a sample of $f_P(p)$ in log-log scale. The approximate straight line in the tail is the signature of a power-law decay of the probability of super-profits. Super-profit outliers are found in the empirical data, although it has not been investigated whether they decay as an approximate power-law.
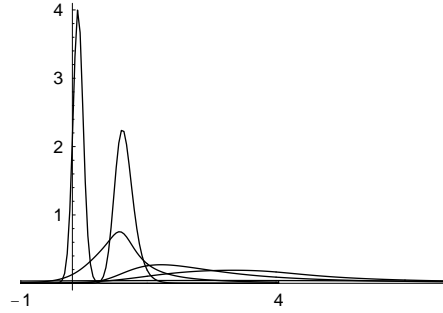
Further analysis of the pdf $f_P(x)$ is required. But the qualitative form of the distribution is sufficiently encouraging to consider it a candidate for fitting to empirical profit measures and for comparison with other candidate functional forms. To go beyond models that assume a realised uniform profit rate it is necessary to investigate empirical data on firm profit and propose theoretical explanations of its distribution. This paper is a tentative step in that direction.

3   Conclusion

A general probabilistic argument suggests that the empirical rate-of-profit distribution will be consistent with a parameter-mix of a ratio of normal variates with means and variances that depend on a firm size parameter that is distributed according to a power law.

(a) Samples plotted on separate scales.

(b) The same samples plotted on a single scale.

Figure 1: Representative numerical samples of the probability density function $f_P(p)$.
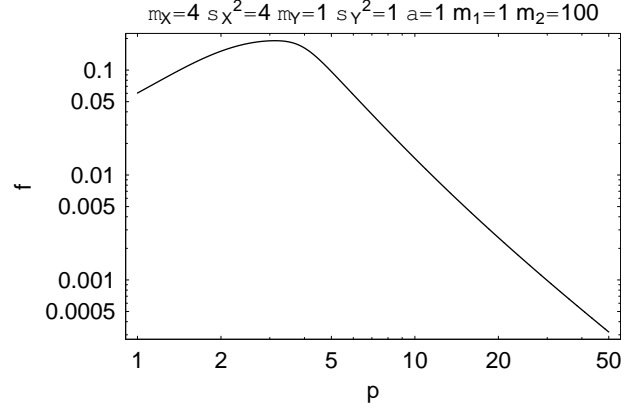
Figure 2: A sample of $f_P(p)$ plotted in log-log scale. Note the long power-law tail.

References

[1] Robert L. Axtell. Zipf distribution of U.S. firm sizes. *Science*, 293:1818–1820, 2001.

[2] Emmanuel Farjoun and Moshe Machover. *Laws of Chaos, a Probabilistic Approach to Political Economy*. Verso, London, 1989.

[3] Yoshi Fujiwara, Corrado Di Guilmi, Hideaki Aoyama, Mauro Gallegati, and Waturu Souma. Do Pareto-Zipf and Gibrat laws hold true? An analysis with European firms. *Physica A*, 335:197–216, 2004.

[4] Edoardo Gaffeo, Mauro Gallegati, and Antonio Palestrini. On the size distribution of firms: additional evidence from the G7 countries. *Physica A*, 324:117–123, 2003.

[5] George Marsaglia. Ratios of normal variables and ratios of sums of uniform variables. *Journal of the American Statistical Association*, 60(309):193–204, 1965.

[6] Julian Wells. What is the distribution of the rate of profit? In *IWGVT mini-conference at the Eastern Economic Association*, New York NY, 2001.

[7] Ian Wright. The social architecture of capitalism. Submitted for publication, preprint at http:// xxx.lanl.gov/ PS_cache/ cond-mat/ pdf/ 0401/ 0401053.pdf, 2004.